

## Tested models for each pupil measure separately

Model 1: Pupil measure ~ SNR/Luminance (lmer, package stats)

Model 2: Pupil measure ~ SNR/Luminance + 1 | Subject (lmer, package lme4) – the one used in the submitted manuscript

Model 3: Pupil measure ~ SNR/Luminance + (1+SNR/Luminance) | Subject (lmer, package lme4)

Model 4: Pupil measure ~ SNR/Luminance + 1 | Subject (lme, package nlme)

Model 5: Pupil measure ~ SNR/Luminance + (1+SNR/Luminance) | Subject (lme, package nlme)

### Selected Models for each pupil measure

*Table 1 Selected Models for each pupil measure separately, italics indicate the old model provided best fit, whereas bold indicates changes in the manuscript and script will be necessary.*

	RPC1	RPC2	RPC3	PPD	MPD	IPA Trial	IPA List
Data set A	<b>Model 2</b>	<b>Model 4</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 1</b>	<i>Model 2</i>	<b>Model 1</b>
Data set B	<b>Model 3</b>	<b>Model 1</b>	<i>Model 1</i>	<i>Model 2</i>	<b>Model 1</b>	<i>Model 2</i>	<b>Model 1</b>

## Principal Component Analysis

### Data set A

#### 1. RPC1 scores

Selected: Model 2 - Residuals within normal distribution (QQplot), narrow confidence intervals. AIC higher than Model 4 of comparable complexity, however still preferable due to the other two criteria.

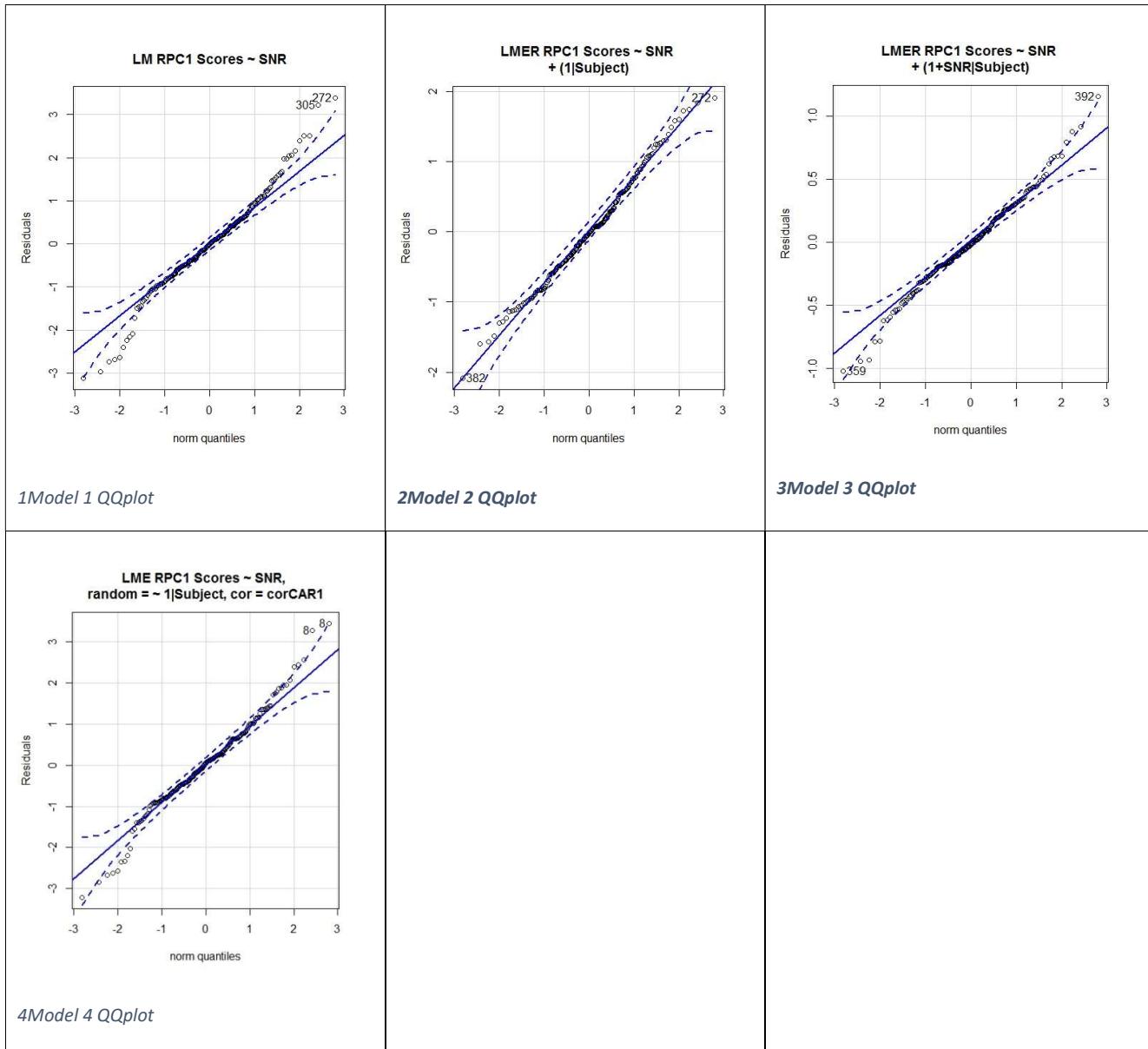
- Convergence

Model 1, Model 2, Model 3, Model 4 – Yes,

Model 5 – No

Further diagnostics for models that converged

- QQplots



- AIC

Model 1 – 586.54, Model 2 – 536.6, **Model 3 – 475.25**, Model 4 – 494.81

- $\beta$  fixed estimates confidence intervals (random effects not included)

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-20dB	<b>1</b>	<b>0.54</b>	<b>1.05</b>	<b>0.51</b>
	2	0.49	1.1	0.61
	3	0.43	1.16	0.73
	4	0.58	1.21	0.63
-10dB	1	-0.32	0.41	0.73
	<b>2</b>	<b>-0.23</b>	<b>0.32</b>	<b>0.55</b>
	3	-0.32	0.41	0.73
	4	-0.4	0.16	0.56

+5dB	<b>1</b>	<b>-1.07</b>	<b>-0.35</b>	<b>0.72</b>
	<b>2</b>	<b>-0.99</b>	<b>-0.44</b>	<b>0.55</b>
	3	-1.13	-0.3	0.83
	4	-1.3	-0.58	0.72

## 2. RPC2 scores

Selected: Model 4 – Narrow estimation intervals, good QQplot, moderate AIC.

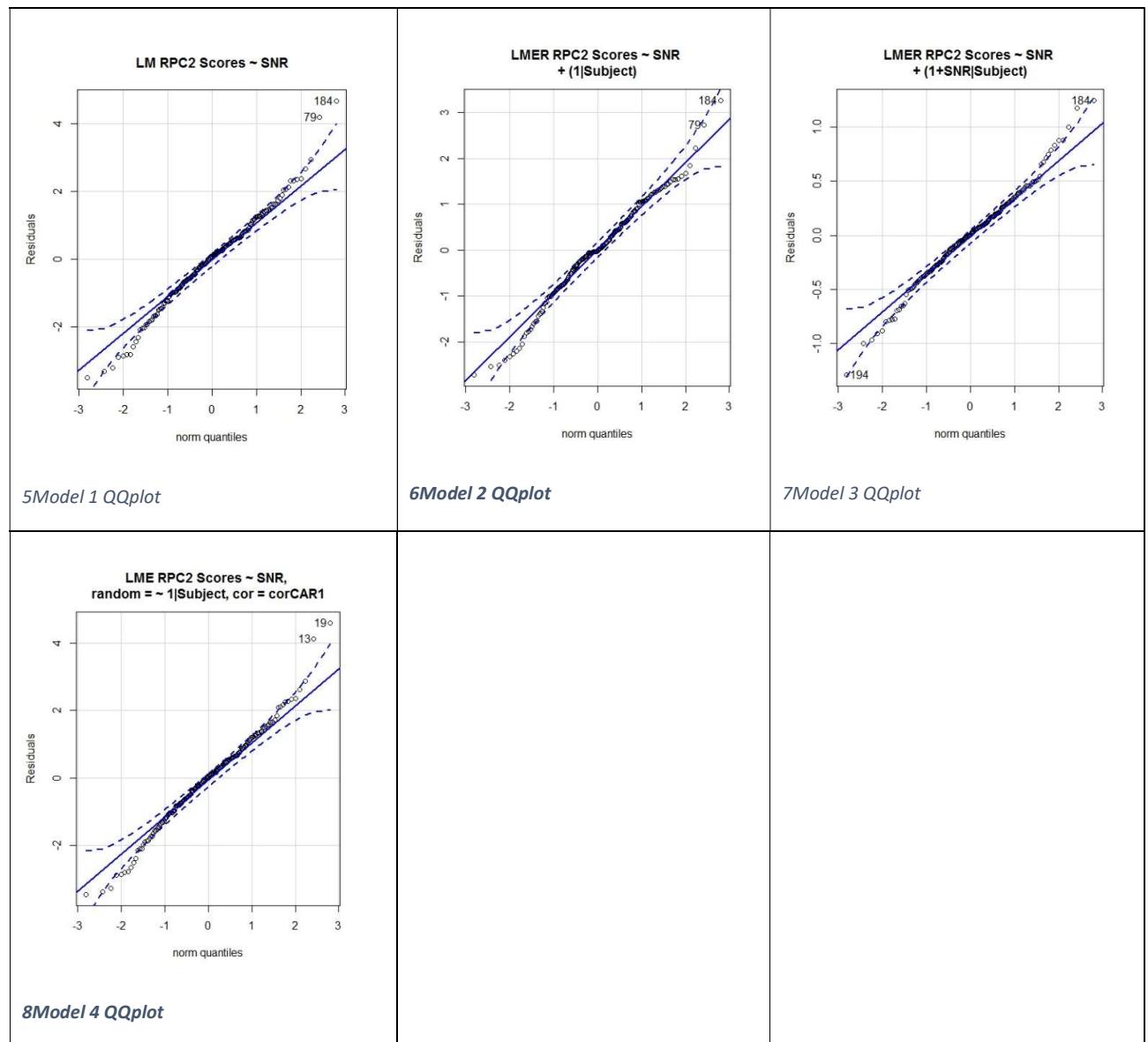
- Convergence

Model 1, 2, 3, 4 – Yes

Model 5 - No

Further diagnostics for models that converged

- QQplots for residuals (both equally good)



- AIC

Model 1 – 670.05, Model 2 – 642.6, **Model 3 – 554.82**, Model 4 – 591.99, Model 5 – 556.66

Anova to choose a better model: L.Ratio(df = 2) 30.05, p < 0.001\*\*\*

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-20dB	<b>1</b>	<b>-1.14</b>	<b>-0.51</b>	<b>0.63</b>
	2	-1.89	-0.47	1.42
	3	-1.26	-0.4	0.86
	4	-1.27	-0.49	0.78
-10dB	1	-0.66	0.23	0.89
	<b>2</b>	<b>-0.58</b>	<b>0.15</b>	<b>0.73</b>
	3	-0.66	0.23	0.89
	<b>4</b>	<b>-0.47</b>	<b>0.26</b>	<b>0.73</b>
+5dB	1	0.97	1.86	0.89
	2	1.04	1.78	0.74
	3	0.79	2.03	1.24
	<b>4</b>	<b>1.08</b>	<b>1.53</b>	<b>0.45</b>

### 3. RPC3 scores

Selected: Model 2 – QQplot slightly worse than for other models, however narrower estimation confidence intervals. AIC comparable with the other Models.

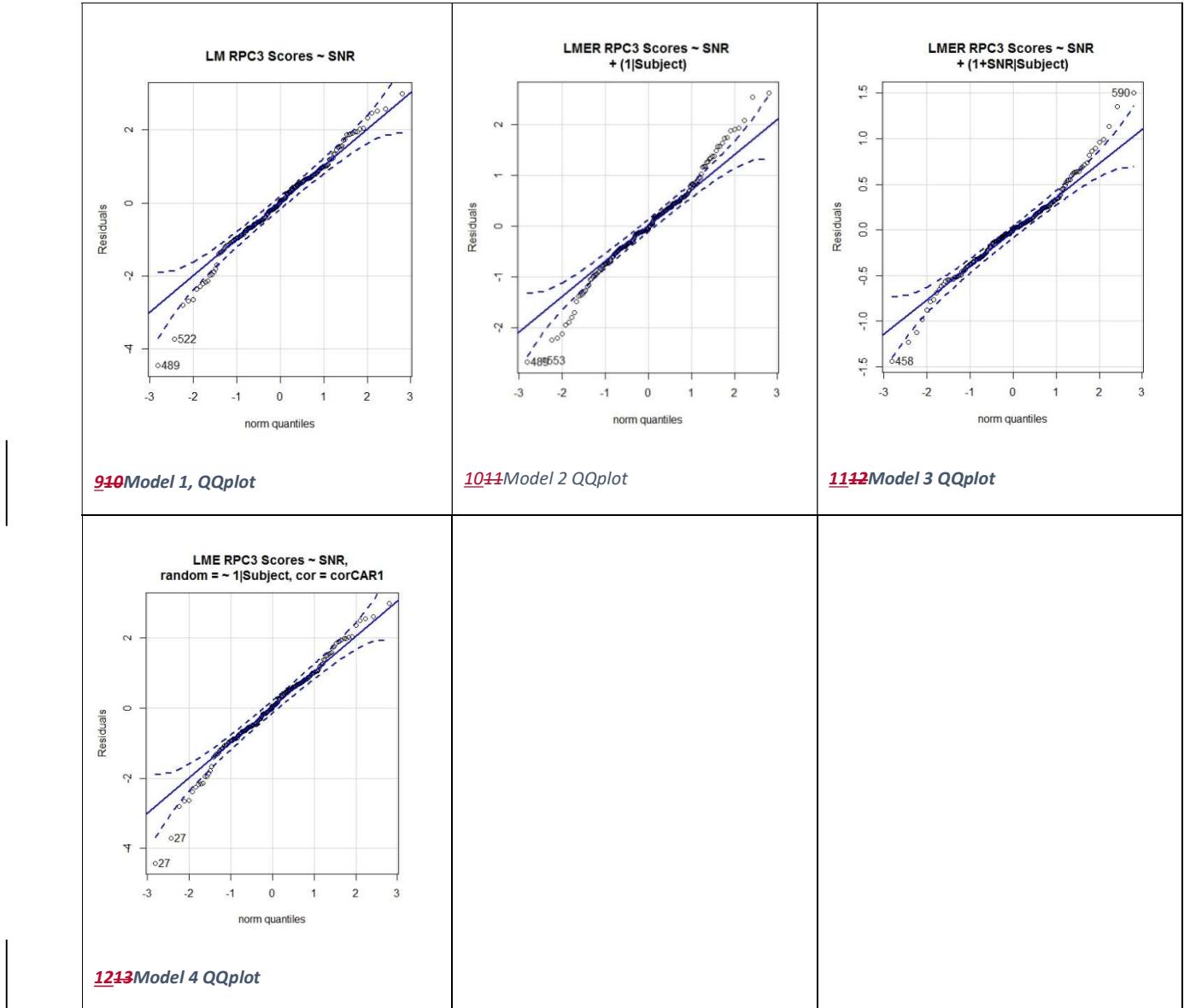
- Convergence

Model 1, Model 2, Model 3, Model 4 – Yes

Model 5 – No

Further diagnostics for models that converged

- QQplots



- AIC

Model 1 – 622.97, Model 2 – 594.37, **Model 3 – 515.56**, Model 4 – 539.24

ANOVA for Model 1 vs. Model 4 L.Ratio(df = 2) = 33.06, p < 0.001\*\*\*

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-20dB	1	<b>0.05</b>	<b>0.61</b>	<b>0.56</b>
	2	0.01	0.65	0.64
	3	-0.02	0.68	0.7
	4	0.01	0.68	0.67
-10dB	1	<b>-0.26</b>	<b>0.53</b>	<b>0.79</b>
	2	-0.19	0.46	0.65
	3	-0.29	0.55	0.84

	<b>4</b>	<b>-0.23</b>	<b>0.41</b>	<b>0.64</b>
+5dB	1	-1.2	-0.41	0.83
	<b>2</b>	<b>-1.13</b>	<b>-0.48</b>	<b>0.65</b>
	3	-1.36	-0.25	1.11
	4	-1.25	-0.45	0.8

## Data set B

### 1. RPC1 scores

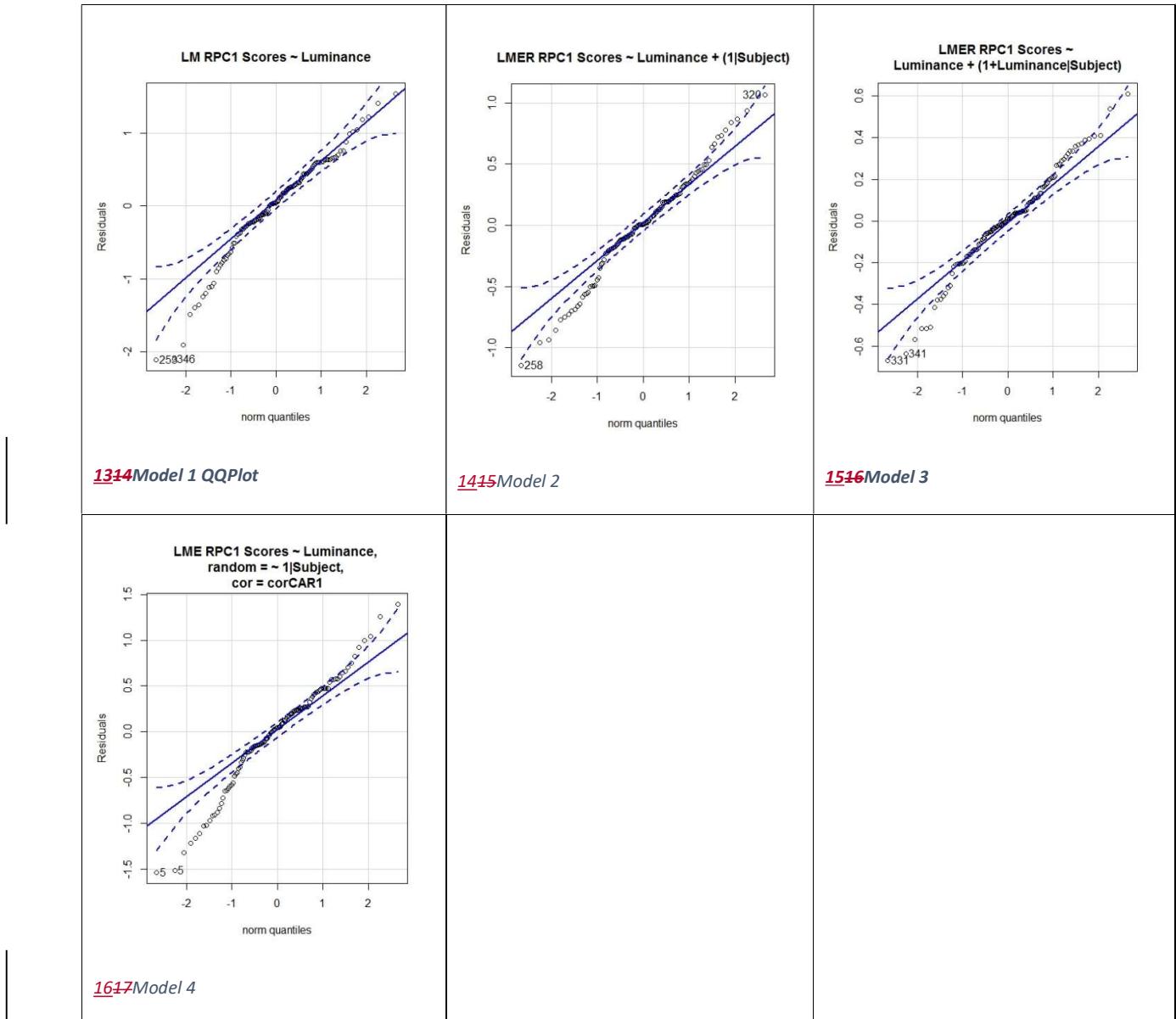
Selected: Model 3 – Best QQplot and balanced confidence intervals. lowest AIC.

- Convergence

Model 1, Model 2, Model 3, Model 4– Yes

Model 5 - No

- QQplots



- AIC

Model 1 – 256.5, Model 2 – 218.5, **Model 3 – 193.6**, Model 4 – 204.5, Model 5 – 195.2

ANOVA on Model 2 and Model 3 Chisq(df = 2) = 28.9, p(>Chisq) < 0.001\*\*\*

ANOVA on Model 4 and Model 5 L.Ratio(df = 3) = 13.4, p = 0.001\*\*

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	<b>1</b>	<b>-0.61</b>	<b>-0.28</b>	<b>0.33</b>
	2	-0.65	-0.24	0.41
	3	-0.65	-0.24	0.41
	4	-0.64	-0.21	0.43
Light	1	-0.43	0.05	0.48

	<b>2</b>	<b>-0.35</b>	<b>-0.03</b>	<b>0.32</b>
	3	-0.43	0.05	0.48
	4	-0.38	-0.02	0.36

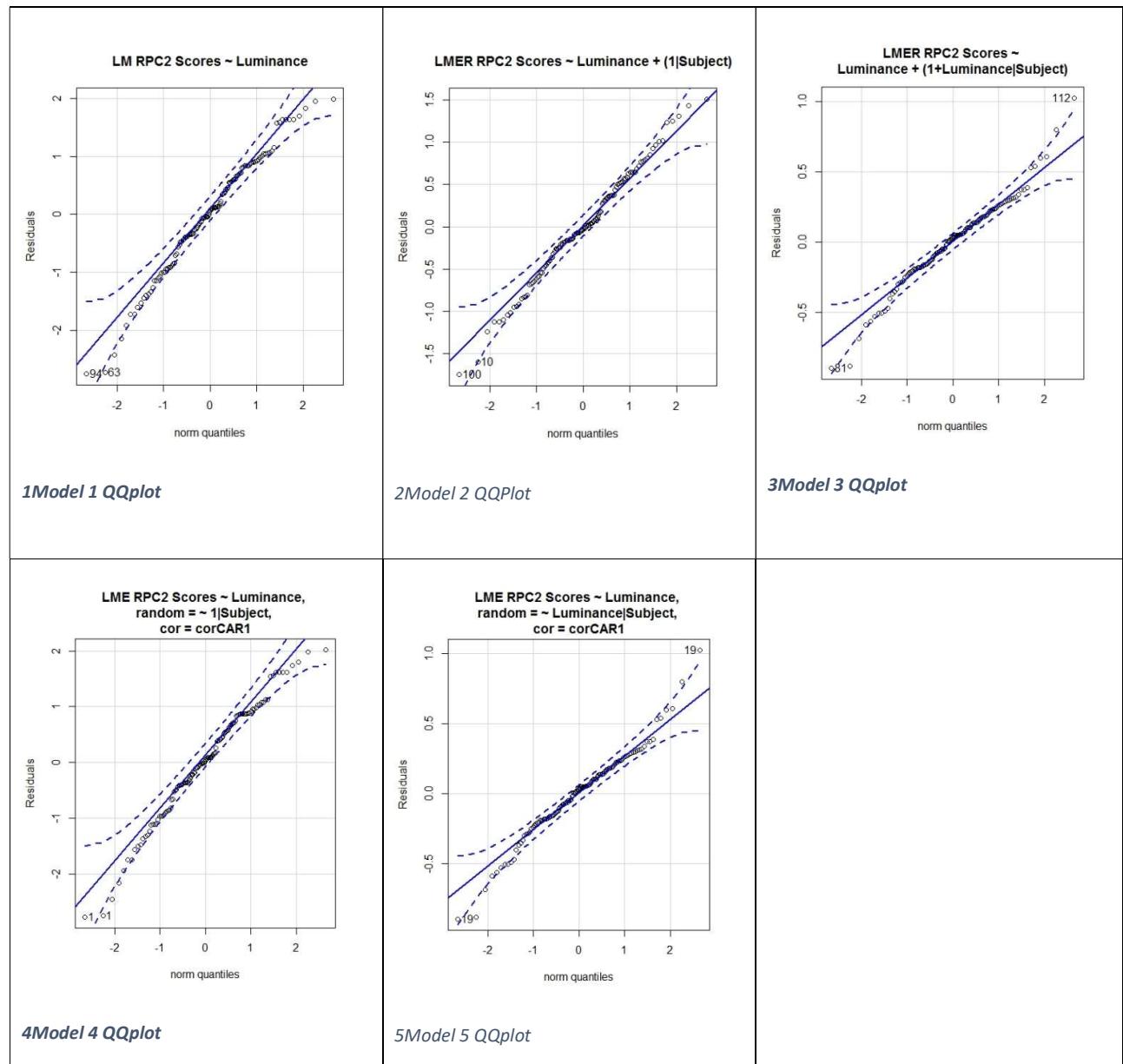
## 2. RPC2 scores

Selected: Model 1 – Both Model 1 and Model 4 present a good fit with QQplot and similar confidence intervals. The simpler model would be preferable.

- Convergence

Model 1, 2, 3 ,4, 5 – Yes

- QQplots



- AIC

Model 1 – 355.42, Model 2 – 319.84, **Model 3 – 279.26**, Model 4 – 290.7, Model 5 – 281.26

ANOVA on Model 4 and Model 1 and Model 5, L.Ratio = 68.72, p < 0.001\*\*\*, L.Ratio = 82.16, p < 0.001\*\*\*. Model 1

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	1	-1.09	-0.59	0.5
	2	<b>-1.05</b>	<b>-0.56</b>	<b>0.49</b>
	3	-1.13	-0.54	0.69
	4	-1.2	-0.54	0.66
	5	<b>-1.13</b>	<b>-0.54</b>	<b>0.49</b>
Light	1	-1.16	-0.45	0.71
	2	-1.05	-0.53	0.52
	3	-1.19	-0.42	0.77
	4	<b>-1</b>	<b>-0.49</b>	<b>0.51</b>
	5	-1.18	-0.42	0.76

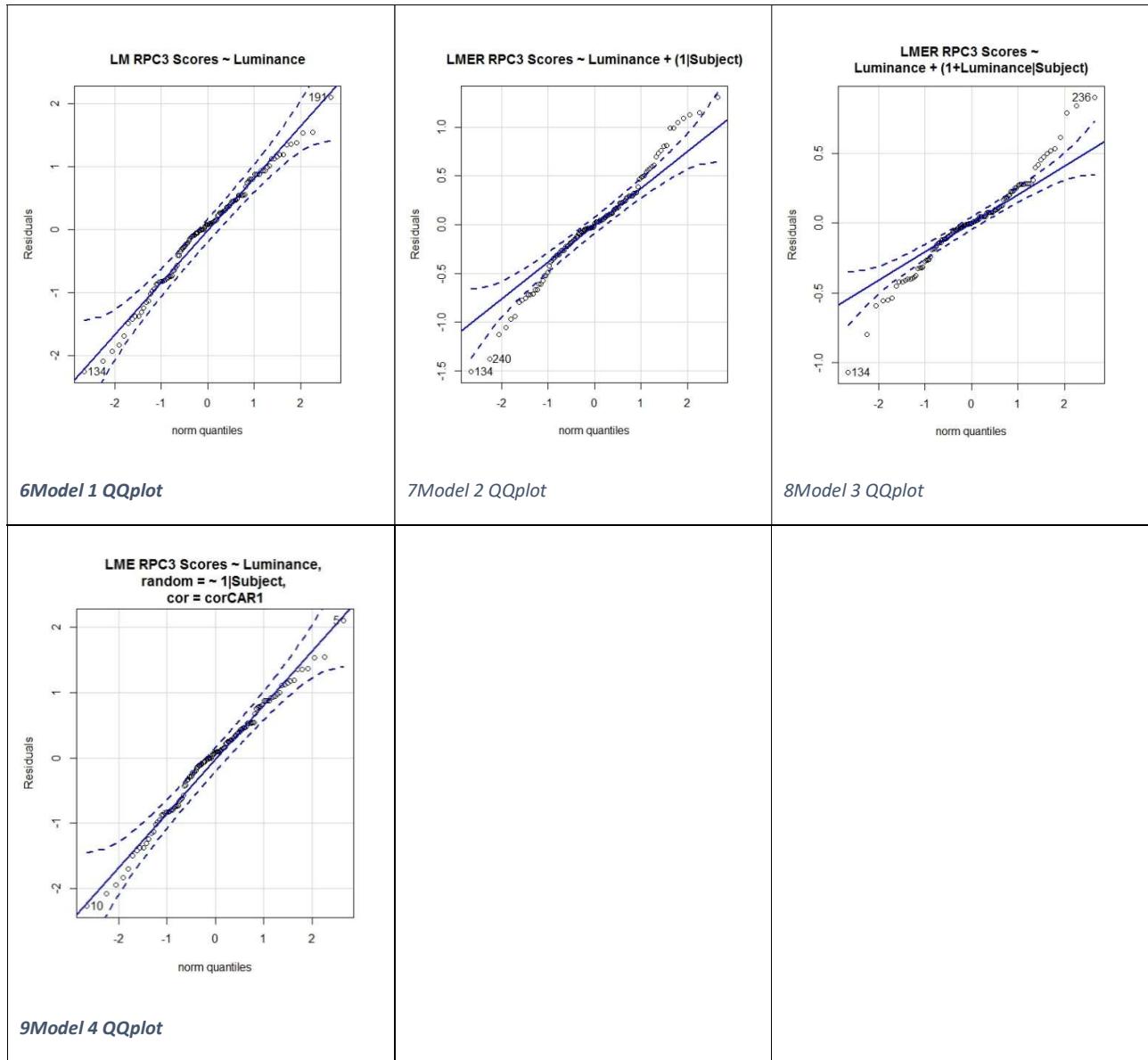
### 3. RPC3 scores

Selected: Model 1 – Both Model 1 and Model 4 present a good fit with QQplot and similar confidence intervals. The simpler model would be preferable.

- Convergence

Model 1, 2, 3, 4 – Yes, 5 - No

- QQplots



- AIC

Model 1 – 308.87, Model 2 – 271.66, **Model 3 – 249.86**, Model 4 – 257.04

ANOVA on Model 1 and Model 2 and Model 3,  $\text{Chisq}(df = 1) = 39.21, p < 0.001^{***}$ ,  $\text{Chisq}(df = 2) = 25.8, p < 0.001^{***}$

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	<b>1</b>	<b>-0.75</b>	<b>-0.33</b>	<b>0.42</b>
	2	-0.8	-0.28	0.52
	3	-0.79	-0.29	0.5
	4	-0.79	-0.27	0.52
Light	1	-0.62	-0.04	0.58
	<b>2</b>	<b>-0.53</b>	<b>-0.13</b>	<b>0.4</b>

	3	-0.63	-0.04	0.59
	4	-0.58	-0.12	0.46

## Index of Pupillary Activity

### Data set A

#### 1. IPA Trial

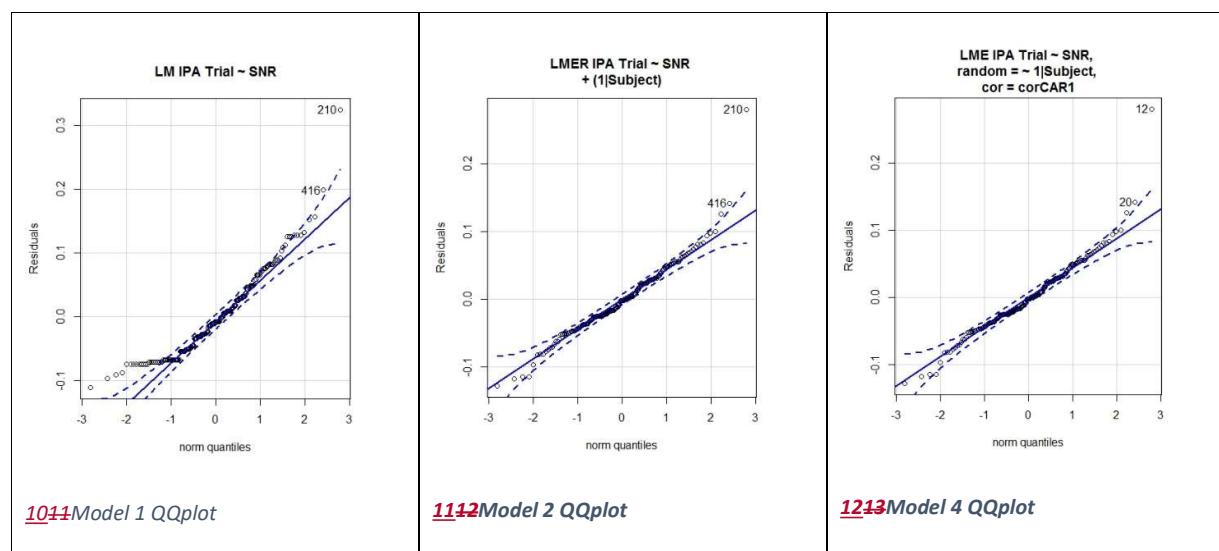
Selected: Model 2 – Models 2 and 4 similarly well fitted to the data. Model 2 has though lower AIC and better handles missing values, that were present in case of this measure. On top of that, auto-correlation term not fully required as IPA is a measure of overall density of dilation reflexes not their timing.

- Convergence

Model 1, Model 2, Model 4 – Yes

Model 3, Model 5 - No

- QQplots



- AIC

Model 1 – -504.74, Model 2 – **-533.51**, Model 4 – -531.51

ANOVA Model 1 and Model 2 Chisq(df = 1) = 30.76, p < 0.001\*\*\*

ANOVA Model 1 and Model 4 L.Ratio = 30.76, p < 0.001\*\*\*

Model 2 and Model 4 very similar fit

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-20dB	1	-0.02	0.03	0.05
	<b>2</b>	<b>-0.01</b>	<b>0.02</b>	<b>0.03</b>

	<b>4</b>	<b>-0.01</b>	<b>0.02</b>	<b>0.03</b>
-10dB	1	0.25	0.28	0.03
	<b>2</b>	<b>0.25</b>	<b>0.29</b>	<b>0.04</b>
	<b>4</b>	<b>0.25</b>	<b>0.29</b>	<b>0.04</b>
+5dB	1	-0.02	0.03	0.05
	<b>2</b>	<b>-0.01</b>	<b>0.02</b>	<b>0.03</b>
	<b>4</b>	<b>-0.01</b>	<b>0.02</b>	<b>0.03</b>

## 2. IPA List

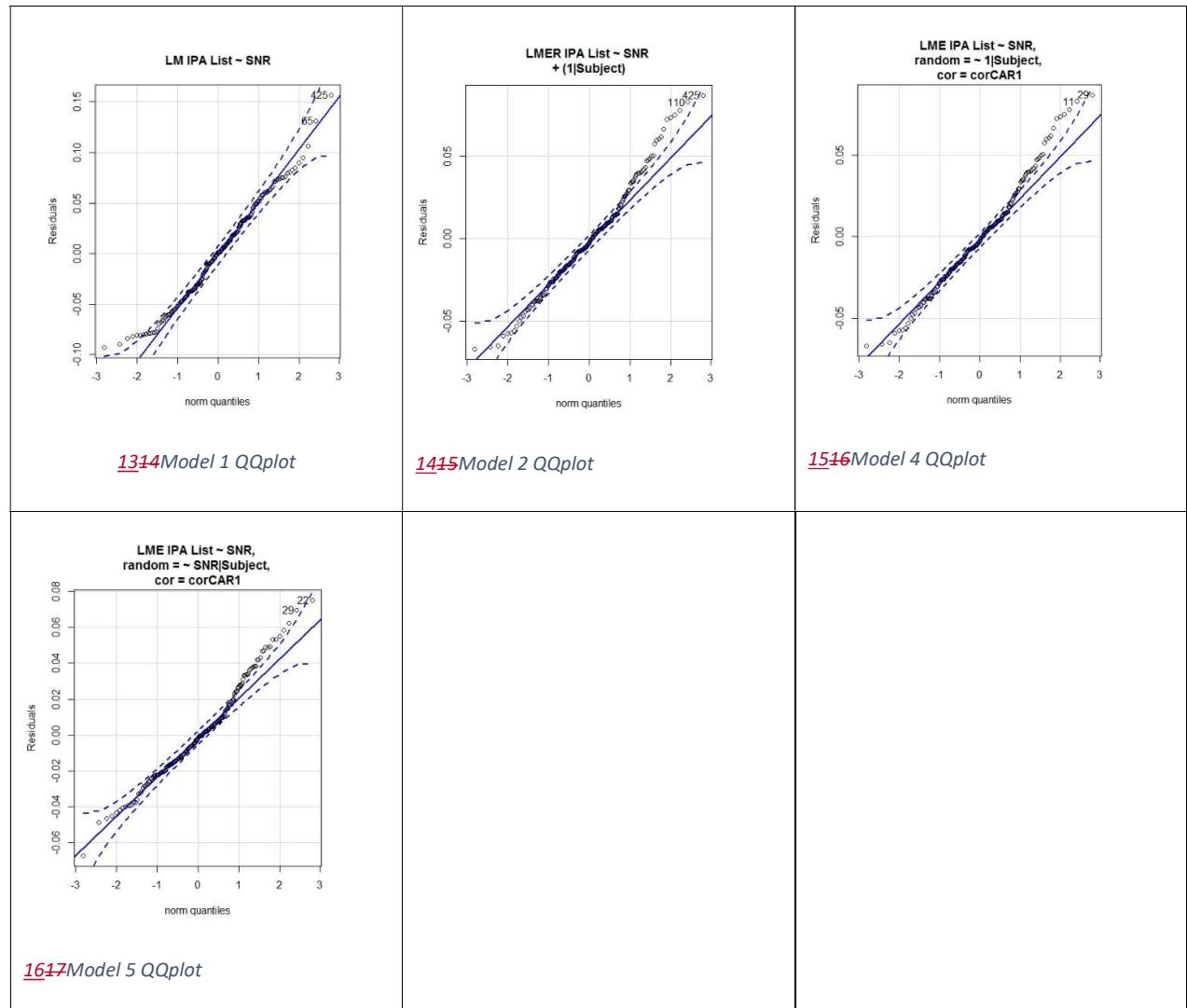
Selected: Model 1 – has best fit according to QQplot and narrow confidence intervals (as all).

- Convergence

Model 1, Model 2, Model 4, Model 5 – Yes

Model 3- No

- QQplots



- AIC

Model 1 – -632.69, Model 2 – -711.32, Model 4 – -709.32, **Model 5 – -713.66**

ANOVA on Model 1 and Model 4 and Model 5 L.Ratio = 80.63, p < 0.001\*\*\*, L.Ratio = 14.34, p = 0.01\*

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-20dB	1	-0.02	0.01	0.03
	2	-0.02	0.02	0.04
	<b>4</b>	<b>-0.01</b>	<b>0.01</b>	<b>0.02</b>
	5	-0.02	0.01	0.03
-10dB	1	0.09	0.12	0.03
	2	0.09	0.12	0.03
	4	0.09	0.12	0.03
	5	0.09	0.12	0.03
+5dB	1	-0.01	0.02	0.03
	2	-0.01	0.02	0.03
	4	-0.01	0.02	0.03
	5	-0.01	0.02	0.03

## Data set B

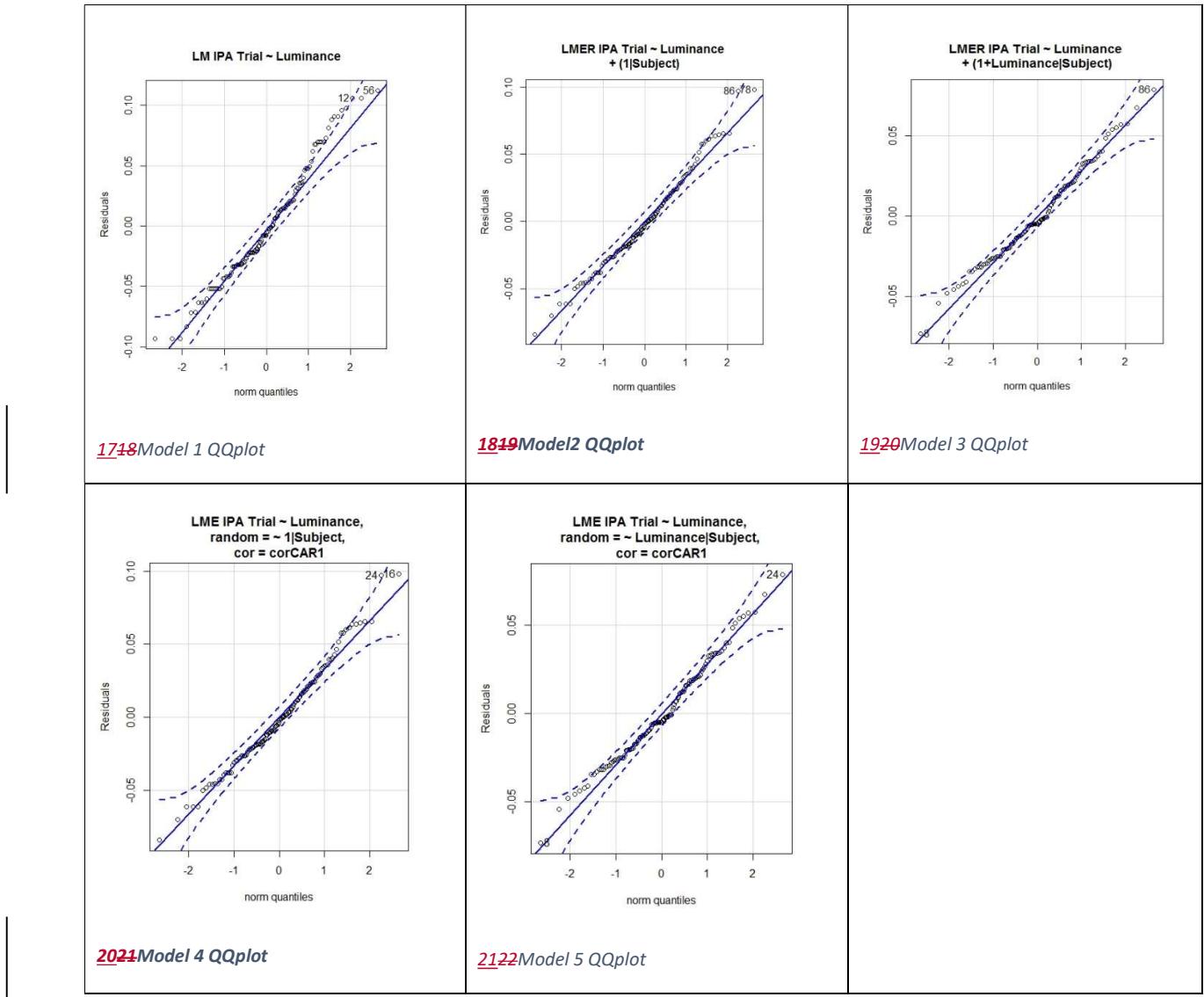
### 1. IPA Trial

Selected: Model 2 – Good fit according to the QQplot, ns difference between Model 2 and Model 3 indicates Model 2 is a better model. In addition Model 2 is better in handling unbalanced data set than Model 4, which was the case here.

- Convergence

Model 1, 2, 3, 4, 5 – Yes

- QQplots



- AIC

Model 1 – -396.31, Model 2 – **-410.76**, Model 3 – **-412.01**, Model 4 – -408.76, Model 5 – -410.01

ANOVA on Model 4 and Model 5 – ns

ANOVA on Model 2 and Model 3 – ns

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	<b>1</b>	<b>0.28</b>	<b>0.3</b>	<b>0.02</b>
	2	0.28	0.31	0.03
	3	0.28	0.31	0.03
	4	0.28	0.31	0.03
	5	0.28	0.31	0.03
Light	1	-0.06	-0.02	0.04
	<b>2</b>	<b>-0.05</b>	<b>-0.03</b>	<b>0.02</b>

	3	-0.06	-0.02	0.04
	<b>4</b>	<b>-0.05</b>	<b>-0.03</b>	<b>0.02</b>
	5	-0.06	-0.02	0.04

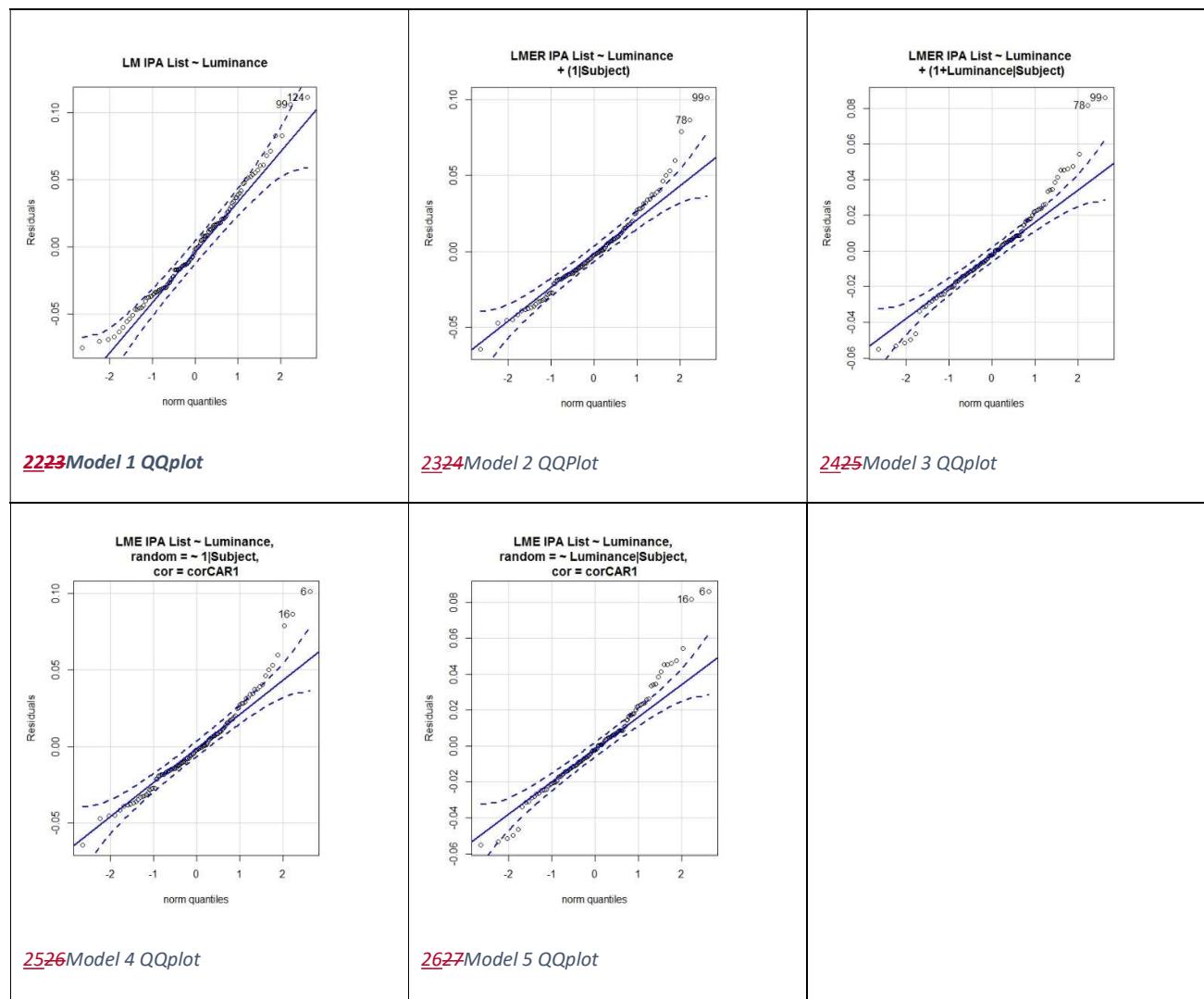
## 2. IPA List

Selected: Model 1 – Both QQplot and confidence intervals show best fit is Model 1. In the QQplots, one can see the prediction of the more complex models become non-normally distributed.

- Convergence

Model 1, 2, 3, 4, 5– Yes

- QQplots



- AIC

Model 1 – -431.94, Model 2 – -447.35, **Model 3 – -453.24**, Model 4 – -445.35, Model 5 – -451.24

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	1	0.09	0.11	0.02
	2	0.09	0.11	0.02
	3	0.09	0.11	0.02
	4	0.09	0.11	0.02
	5	0.09	0.11	0.02
Light	1	-0.03	-0.01	0.02
	2	-0.03	-0.01	0.02
	3	-0.03	-0.01	0.02
	4	-0.03	-0.01	0.02
	5	-0.03	-0.01	0.02

## Mean Pupil Dilation

### Data set A

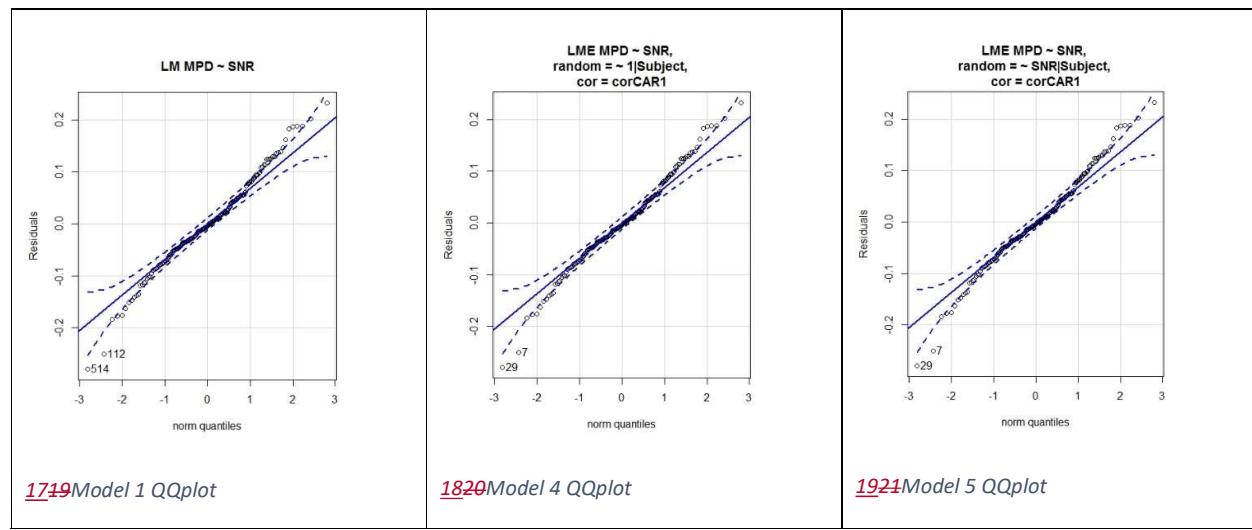
Selected: Model 1 – Lowest AIC, QQplots similar to all models, so simpler one would be in favour.  
 Confidence intervals a bit broader, however they are similar to all, simpler model would be in favour.

- Convergence

Model 1, Model 4, Model 5 – Yes

Model 2, Model 3 - No

- QQplots



- AIC

**Model 1 – -425.11, Model 4 – -421.11 , Model 5 – -411.11**

ANOVA Model 1 and Model 4 and Model 5 - ns

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-10dB	1	0.05	0.09	0.04

	<b>4</b>	<b>0.05</b>	<b>0.07</b>	<b>0.02</b>
	5	0.05	0.09	0.04
-20dB	<b>1</b>	<b>-0.04</b>	<b>0.02</b>	<b>0.06</b>
	<b>4</b>	<b>-0.04</b>	<b>0.02</b>	<b>0.06</b>
	<b>5</b>	<b>-0.04</b>	<b>0.02</b>	<b>0.06</b>
+5dB	1	-0.12	-0.06	0.06
	<b>4</b>	<b>-0.12</b>	<b>-0.09</b>	<b>0.03</b>
	5	-0.12	-0.06	0.06

### Data set B

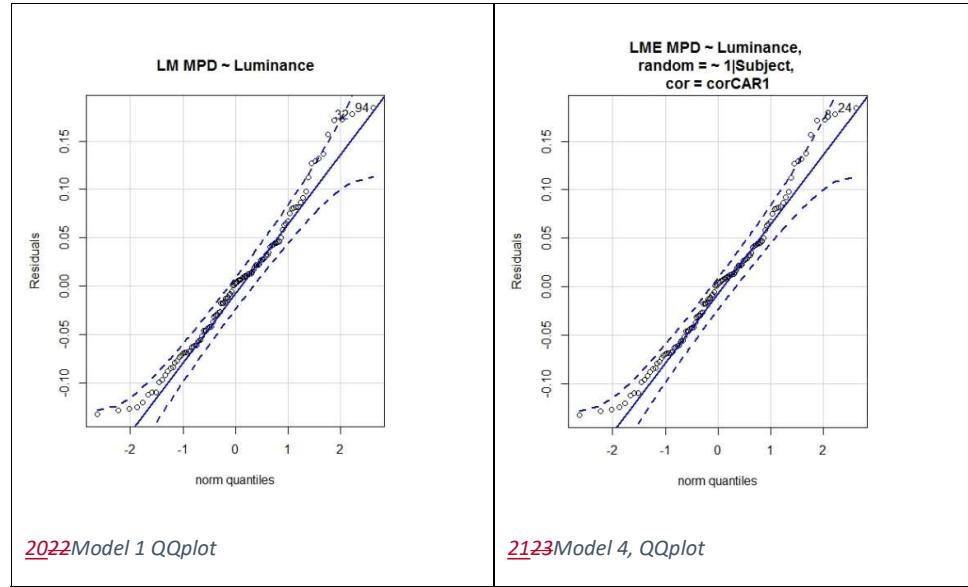
Selected: Model 1 – Lower AIC, no difference in confidence intervals or QQplots, simpler model would be in favour

- Convergence

Model 1, Model 4 – Yes

Model 2, Model 3, Model 5 - No

- QQplots



- AIC

**Model 1 – -276.89, Model 4 – -273.13**

ANOVA for Model 4 and Model 1 – ns.

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	1	0.05	0.09	0.04
	4	0.05	0.09	0.04
Light	1	0.03	0.08	0.05
	4	0.03	0.08	0.05

## Peak Pupil Dilation

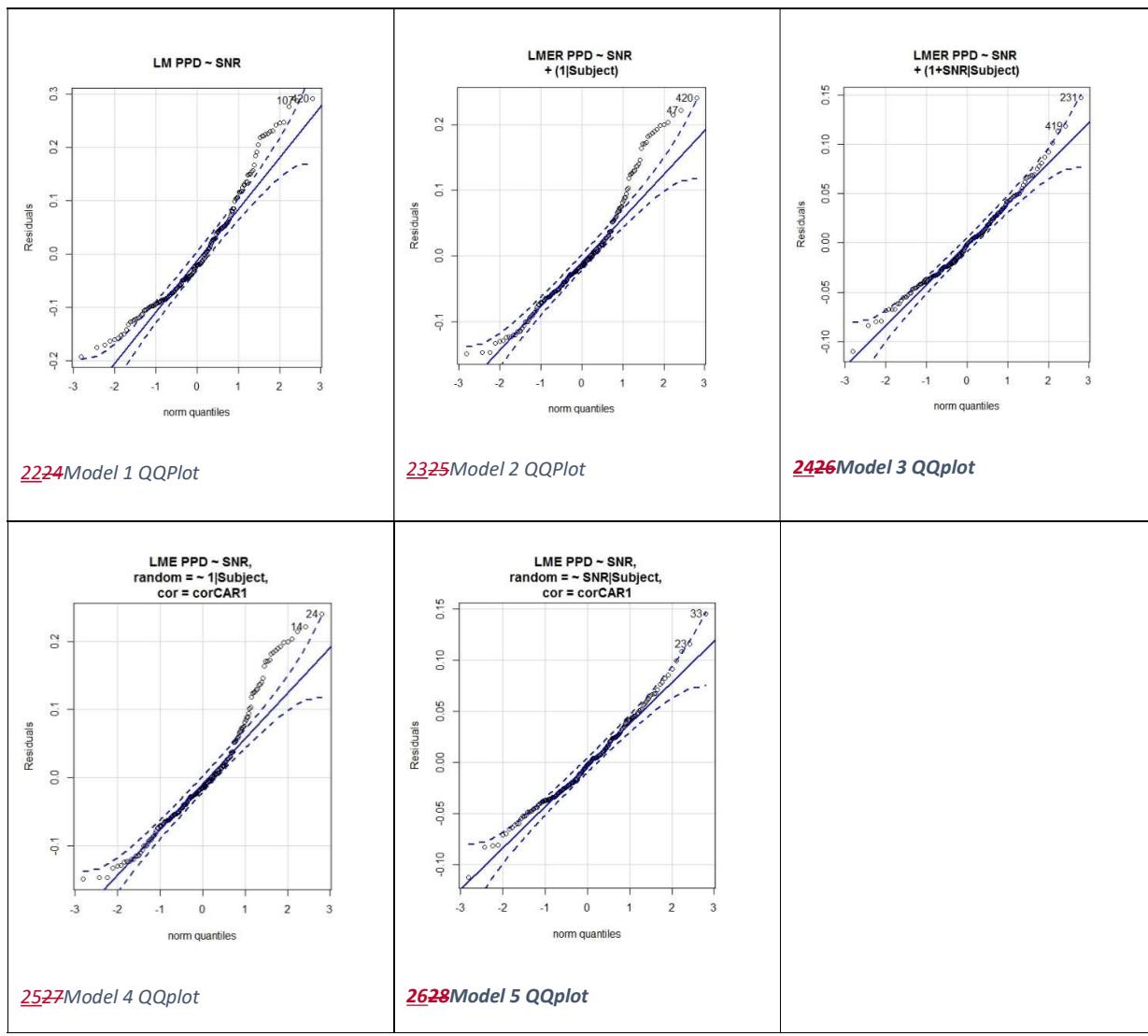
### Data set A

Selected: Model 3 – Variance is not the same within SNR conditions, so it makes sense to account for that in the model. Auto-correlation maybe not crucial in case of PPD

- Convergence

Model 1, 2, 3, 4, 5 – Yes

- QQplots



- AIC

Model 1 – -327.95, Model 2 – -349.25, **Model 3 – -415.38**, Model 4 – -347.25, Model 5 – -414.25

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
-10dB	<b>1</b>	<b>0.19</b>	<b>0.24</b>	<b>0.05</b>
	2	0.18	0.24	0.06
	3	0.17	0.25	0.08
	4	0.18	0.24	0.06
	5	0.17	0.25	0.08
-20dB	1	-0.08	-0.006	0.074
	<b>2</b>	<b>-0.07</b>	<b>-0.01</b>	<b>0.06</b>
	3	-0.09	0.006	0.096
	<b>4</b>	<b>-0.07</b>	<b>-0.01</b>	<b>0.06</b>
	5	-0.09	0.005	0.085
+5dB	1	-0.13	-0.05	0.08
	<b>2</b>	<b>-0.12</b>	<b>-0.06</b>	<b>0.06</b>
	3	-0.13	-0.05	0.08
	<b>4</b>	<b>-0.12</b>	<b>-0.06</b>	<b>0.06</b>
	5	-0.13	-0.05	0.08

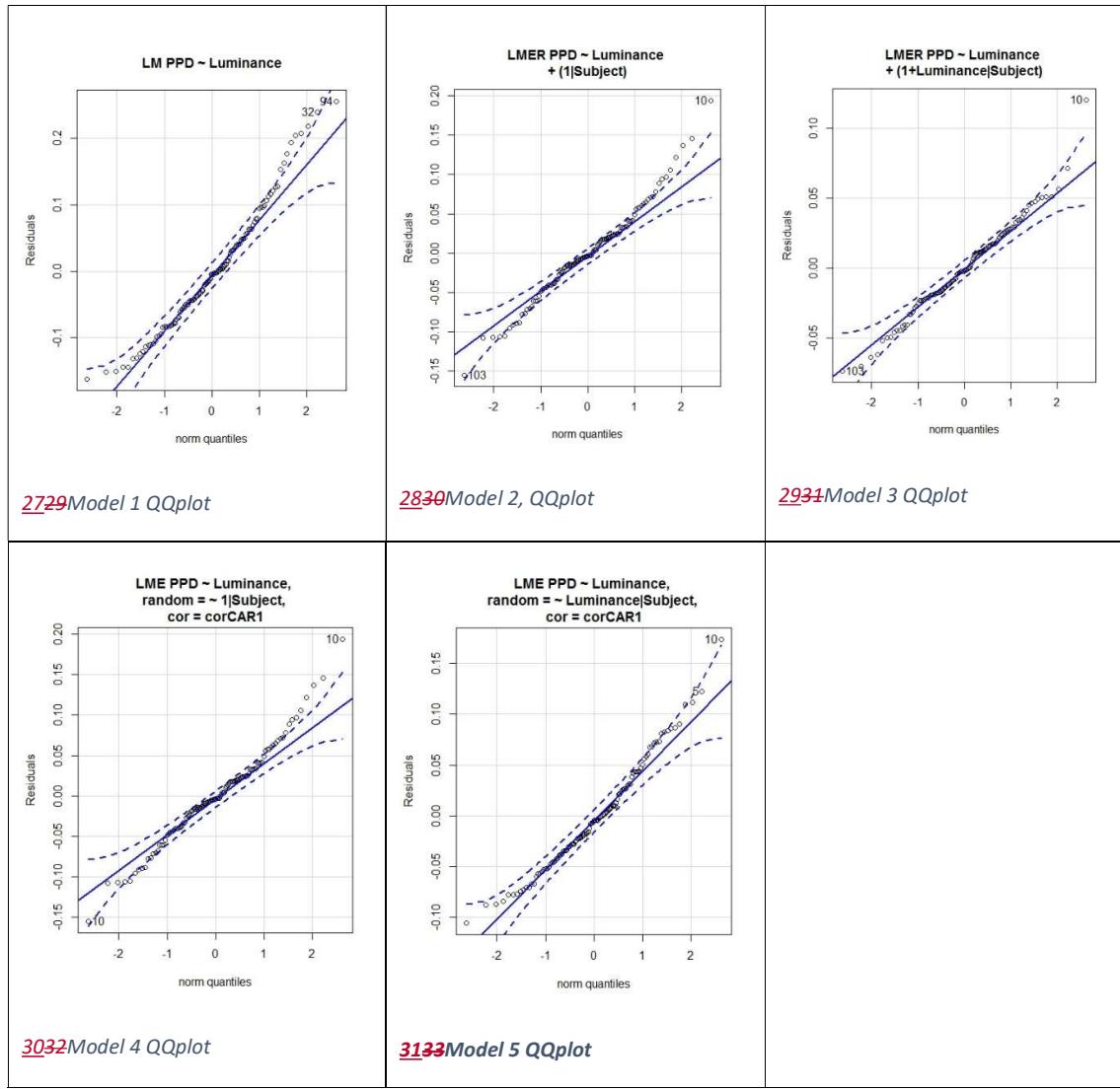
### Data set B

Selected: Model 2 – 5 with auto-correlation term is maybe not that important in case of PPD as this pupil measure should be immune to time. Model 2 does have narrow confidence intervals in both conditions.

- Convergence

Model 1, 2, 3, 4, 5 – Yes

- QQplots



- AIC

Model 1 – -222.45, Model 2 – -256.57, Model 3 – -286.71, Model 4 – -254.57, **Model 5 – -292.11**

- $\beta$  estimates confidence intervals

Condition	Model	Lower (2.5%)	Upper (97.5%)	dConf
Dark	1	0.14	0.19	0.05
	<b>2</b>	<b>0.13</b>	<b>0.09</b>	<b>0.04</b>
	3	0.14	0.19	0.05
	4	0.13	0.19	0.06
	5	0.14	0.19	0.05
Light	1	0.036	0.1	0.064
	2	0.05	0.09	0.04
	3	0.03	0.1	0.07
	4	0.05	0.09	0.04
	<b>5</b>	<b>0.04</b>	<b>0.07</b>	<b>0.03</b>